

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
"КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
ІМЕНІ ІГОРЯ СІКОРСЬКОГО"**

Ільїн Микола Іванович



УДК 681.513.675

**Інформаційна технологія екологічного моніторингу процесів
забруднення атмосфери з використанням високопродуктивних
обчислень**

Спеціальність: 05.13.06 – інформаційні технології

Автореферат
дисертації на здобуття наукового ступеня
кандидата технічних наук

Київ – 2019

Дисертацією є рукопис.

Робота виконана на кафедрі інформаційної безпеки Фізико-технічного інституту Національного технічного університету України "Київський політехнічний інститут імені Ігоря Сікорського" Міністерства освіти і науки України.

Науковий керівник: доктор технічних наук, професор
Новіков Олексій Миколайович,
Національний технічний університет України
"Київський політехнічний інститут імені Ігоря
Сікорського", проректор з науково-педагогічної роботи
(перспективний розвиток)

Офіційні опоненти: доктор технічних наук, доцент
Машталір Сергій Володимирович
Харківський національний університет
радіоелектроніки, професор кафедри інформатики

доктор технічних наук, старший науковий співробітник
Іванюта Сергій Петрович,
Національний інститут стратегічних досліджень,
заступник завідувача відділу енергетичної та
техногенної безпеки.

Захист відбудеться 27 червня 2019 р. об 11.00 годині на засіданні спеціалізованої вченої ради Д 26.002.29 у Національному технічному університеті України "Київський політехнічний інститут імені Ігоря Сікорського" за адресою: 03056, м. Київ, пр. Перемоги 37, корпус №11, аудиторія № 215.

З дисертацією можна ознайомитись у науково-технічній бібліотеці Національного технічного університету України "Київський політехнічний інститут імені Ігоря Сікорського" за адресою: 03056, м. Київ, пр. Перемоги, 37.

Автореферат розісланий 27 травня 2019 р.

Вчений секретар
спеціалізованої вченої ради Д 26.002.29
доктор технічних наук, професор

С.Ф.Теленик

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми. Інформаційні технології є основою розвитку багатьох напрямків наукового пізнання, сучасних техніки та технологій. Основними напрямками інформаційних технологій, які зараз знаходяться в бурхливому розвитку, є інформаційно-комунікаційні технології, інтелектуальні методи, оброблення великих даних, високопродуктивні обчислювальні технології. Серед галузей знань, де визначальним є використання сучасних високопродуктивних обчислювальних технологій, є дослідження складних фізичних процесів і полів, серед яких екологічні процеси забруднення атмосфери, ґрунтових вод, сейсмічної активності, аналіз даних дистанційного зондування Землі та ін. В наведених галузях переважна більшість задач мають трансобчислювальну складність.

Одним з основних підходів до подолання трансобчислювальної складності є розроблення та використання інформаційних технологій організації обчислень за рахунок поєднання спеціалізованих технічних засобів та математичних методів. В якості технічних засобів використовується паралельні високопродуктивні обчислювальні комплекси, зокрема високопродуктивні паралельні гібридні системи з використанням спеціалізованих прискорювачів. Математичні методи базуються на декомпозиції вихідної задачі на окремі складові з метою їх паралельного вирішення на окремих процесорах з наступним агрегуванням отриманих результатів, розробленні та використанні спеціалізованих чисельних методів з урахуванням обмежень існуючих високопродуктивних обчислювальних систем та ін. При масовому застосуванні цих методів для задач довгострокового моніторингу важливим фактором є також вартість обчислень – енергоефективність та вартість застосованих обчислювальних комплексів. На сьогодні найбільш ефективними за сукупними показниками є гібридні системи на базі графічних процесорів архітектури Nvidia CUDA. Продуктивність таких систем на 1-3 порядки вища ніж CPU і на 1-2 порядки споживають менше енергії для обчислення однієї операції з плаваючою точкою (FLOP). Разом з тим такі системи мають суттєві обмеження на обсяг оперативної пам'яті на обчислювальне ядро – на 3-4 порядки менше ніж у центрального процесора традиційного кластера. Тому розроблення спеціальних математичних методів з урахуванням технічних обмежень на обсяг пам'яті, енергоефективність, високий ступінь паралелізму, як складової інформаційної технології моніторингу, є актуальною задачею.

Розроблення інформаційних технологій екологічного моніторингу процесів забруднення атмосфери вирішувалися в роботах З. Златева, І. Дімова, М. З. Згуровського, І. В. Сергієнка, В. В. Скопечького, М. З. Железняка, І. В. Ковальця, А. Б. Качинського та ін. Разом з тим прогрес апаратного забезпечення обчислювальних систем, розвиток апаратних прискорювачів на базі графічних процесорів загального призначення робить актуальними розроблення нових інформаційних технологій екологічного моніторингу, спеціалізованих математичних методів та комп'ютерних систем, які найбільш ефективно використовують зазначені вище можливості.

Зв'язок роботи з науковими програмами, планами, темами. Робота виконана на кафедрі інформаційної безпеки Національного технічного університету

України "Київський політехнічний інститут імені Ігоря Сікорського" Міністерства освіти і науки України в рамках державних бюджетних тем: «Теоретичні та прикладні задачі безпеки структурно-складних систем» (номер державної реєстрації 0110U000489); «Логіко-ймовірнісний підхід в задачах безпеки структурно-складних систем» (номер державної реєстрації 0113U002468).

Мета і задачі дисертаційного дослідження. Метою дисертаційного дослідження є підвищення оперативності, точності та мобільності систем екологічного моніторингу процесів забруднення атмосфери за рахунок розроблення та застосування інформаційної технології з використанням високопродуктивних обчислень, яка базується на спеціалізованих математичних методах та високопродуктивних графічних прискорювачах загального призначення.

Об'єктом дослідження є процеси оброблення інформації при моделюванні процесів забруднення атмосфери хімічно-взаємодіючими домішками.

Предметом дослідження є методи моделювання, паралельні алгоритми, методи параметричної ідентифікації та інформаційні технології моніторингу процесів забруднення атмосфери хімічно-взаємодіючими домішками з урахуванням аномальних властивостей.

Основні задачі дослідження, у відповідності до поставленої мети, полягають у наступному:

1. Проаналізувати математичні методи моніторингу процесів забруднення атмосфери, як складової інформаційних технологій з використанням високопродуктивних паралельних гібридних обчислювальних систем з використанням графічних прискорювачів;

2. Удосконалити існуючі моделі та паралельні обчислювальні методи моделювання процесів забруднення атмосфери хімічно-взаємодіючими домішками з урахуванням аномальних властивостей самоочищення;

3. Удосконалити методи параметричної ідентифікації функцій джерел забруднюючих домішок та параметрів самоочищення процесів атмосферної дифузії на основі розроблених паралельних методів вирішення систем лінійних алгебраїчних рівнянь (далі СЛАР) великої розмірності;

4. Розробити інформаційну технологію екологічного моніторингу процесів забруднення хімічно-взаємодіючими домішками з урахуванням аномальних властивостей самоочищення, яка базується на застосуванні гібридних обчислювальних систем з використанням графічних прискорювачів.

5. Провести експериментальне дослідження розроблених методів та інформаційної технології екологічного моніторингу.

Методи дослідження. В роботі використовувались методи математичної фізики, теорія варіаційних нерівностей, методи параметричної ідентифікації фізичних процесів, метод функціональної параметризації та методи організації паралельних обчислень.

Наукова новизна одержаних результатів визначається наступними положеннями:

1. Вперше запропоновано для моделювання процесів забруднення атмосфери використовувати комп'ютерну систему на базі графічних прискорювачів архітектури Nvidia CUDA, як складову розробленої інформаційної технології, яка

характеризується підвищеною швидкістю, низькою вартістю елементної бази, на відміну від суперкомп'ютерних технологій, але має низькі ресурси оперативної пам'яті, що обмежує використання традиційних методів обчислювальної математики.

2. Вперше розроблено паралельні прямий та гібридний методи вирішення СЛАР великої розмірності із специфічними розрідженими матрицями, які є скінченно-вимірним представленням моделі процесів забруднення атмосфери та надають можливість ефективного застосування гібридних обчислювальних систем на базі графічних прискорювачів архітектури Nvidia CUDA.

3. Удосконалено метод параметричної ідентифікації функцій джерел моделі процесів забруднення атмосфери хімічно-взаємодіючими домішками з врахуванням аномальних ефектів самоочищення, що характеризується підвищеною точністю моделювання та швидкістю за рахунок використання розроблених паралельних прямого та гібридного методів вирішення СЛАР великої розмірності.

4. Вперше розроблено інформаційну технологію моніторингу аномальних процесів забруднення атмосфери, яка складається з поєднання технологій збору, оброблення, накопичення та надання кінцевому користувачу результатів моніторингу на основі методів реалізації моделей забруднення атмосфери, що уможливають використання гібридних високопродуктивних обчислювальних систем архітектури Nvidia CUDA, і забезпечують підвищення оперативності і точності моніторингу та розширення області його охоплення за рахунок можливості використання мобільних комплексів.

Практичне значення одержаних результатів. На основі розроблених методів, моделей та алгоритмів розроблена і реалізована інформаційна технологія, яка є основою для побудови спеціалізованого програмно-апаратного комплексу «Моніторинг».

Результати дисертаційної роботи використовуються в ТОВ «Укркотехінвест» і навчальному процесі кафедр інженерної екології Інституту енергозбереження та енергоменеджменту та інформаційної безпеки Фізико-технічного інституту Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського».

Особистий внесок здобувача. Здобувачем особисто вдосконалено метод параметричної ідентифікації функцій джерел забруднення в моделях аномальних процесів забруднення атмосфери, розроблено метод паралельної обчислювальної реалізації моделей та інформаційну технологію моніторингу аномальних процесів забруднення атмосфери хімічно-взаємодіючими домішками.

Апробація результатів дисертації. Основні положення і результати дисертаційної роботи були представлені, обговорювалися та доповідались на 3-х міжнародних конференціях:

- Міжнародна наукова конференція студентів, аспірантів і молодих вчених «Теоретичні і прикладні аспекти кібернетики» (м. Київ, 2011)
- VI міжнародна конференція «Екологічна безпека: проблеми і шляхи вирішення» (м. Харків, 2010 р.)
- Міжнародна науково-технічна конференція «Системний аналіз та інформаційні технології» (м. Київ, 2011 р.)

Публікації. За матеріалами дисертаційного дослідження опубліковано 9 наукових праць, з них 6 статей у наукових фахових виданнях (у тому числі 3 у виданнях України, які включені до міжнародних наукометричних баз) та 3-х тезах доповідей в збірниках матеріалів конференцій.

Структура і обсяг дисертації. Дисертація складається зі вступу, чотирьох розділів, висновків, списку використаних джерел з 162 найменувань. Загальний обсяг дисертації 243 сторінок, в тому числі 146 сторінок основного тексту та 18 сторінок використаних джерел. Робота містить 16 рисунків, 13 таблиць

ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** обґрунтовано актуальність теми дисертаційної роботи, сформульовано мету і задачі дослідження, наукову новизну та практичне значення отриманих результатів. Наведено дані про впровадження результатів роботи, її апробацію, публікації та особистий внесок здобувача.

У **першому розділі** виконано аналіз сучасного стану моделювання процесів розповсюдження забруднюючих домішок в атмосфері. Виконано огляд існуючих моделей та методів дослідження процесів атмосферної дифузії. Проаналізовано рекомендації зі застосування моделей розповсюдження шкідливих домішок Агентства з охорони навколишнього середовища США (Environment Protection Agency Support Center for Regulatory Atmospheric Modeling, Air Quality Modeling Group) та Європейського агентства навколишнього середовища (European Environment Agency Model Documentation System). На основі проведеного аналізу сформульовано актуальність теми дисертаційного дослідження, яка обумовлена обчислювальною складністю Ейлеревих моделей розповсюдження шкідливих домішок в атмосфері та задач ідентифікації одностороннього процесу атмосферної дифузії при обмеженні на граничну концентрацію забруднюючої домішки.

Наведено огляд високопродуктивних інформаційних технологій реалізації та паралельних алгоритмів для розв'язання рівнянь в частинних похідних, що виникають при математичній формалізації моделей атмосферної дифузії. Показано актуальність створення нових інформаційних технологій екологічного моніторингу, спеціалізованих математичних методів та комп'ютерних систем, які найбільш ефективно використовують паралельні алгоритми СЛАР та відповідні інформаційні технології на базі кластерів графічних процесорів архітектури Nvidia CUDA.

Другий розділ присвячено вирішенню задачі підвищення обчислювальної ефективності моделювання процесів забруднення атмосфери хімічно-взаємодіючими домішками з урахуванням аномальних властивостей.

Процес розповсюдження забруднюючих домішок в атмосфері з урахуванням ефекту самоочищення при перевищенні граничної концентрації забруднюючої домішки може бути представлений у формі системи варіаційних нерівностей вигляду:

$$\left(\frac{\partial u_s}{\partial t}, v_s - u_s \right) + (A(\lambda)u_s, v_s - u_s) + \psi_s(v_s) - \psi_s(u_s) \geq (f_s, v_s - u_s), \quad s = 1, \dots, N_s \quad (1)$$

$$\text{в } Q \forall v \in H^1(\Omega) = V \quad (2)$$

з початковими умовами

$$u_s|_{t=0} = u_{s0} \text{ в } \Omega, \quad s=1, \dots, N_s, \quad (3)$$

де $u_s(t, z)$ – функція концентрації забруднення визначена на обмеженій відкритій множині Ω простору \mathbb{R}^n , $n=1, 2, 3$, з гладкою границею Γ на часовому інтервалі $(0, t_k)$ для $t_k < \infty$, $\Theta = \Omega \times (0, t_k)$, $s=1, \dots, N_s$, N_s – кількість домішок, (f, g) – дія функціоналу $f \in (H^1(\Omega))^*$ на елемент $g \in H^1(\Omega)$, $A(\lambda): V \rightarrow V^*$ – оператор, визначений білінійною формою;

$$\begin{aligned} (A(\lambda)u_s, \xi) = & \sum_{i=1}^n \int_{\Omega} \left(k_i(z) \frac{\partial u_s}{\partial z_i} \frac{\partial \xi}{\partial z_i} - c_i(z) \frac{\partial u_s}{\partial z_i} \xi \right) dz + \\ & + \int_{\Omega} d(z) u_s \xi dz, \quad \forall \xi \in H^1(\Omega) = V, \quad s=1, \dots, N_s, \end{aligned}$$

де $d(z)$ враховує нелінійний характер хімічної взаємодії забруднюючих домішок і має вигляд

$$d(z) = D(z) + Q_s(u_1 \dots u_{N_s}), \quad s=1, \dots, N_s,$$

де $D(z)$ – процеси типу “джерело-сток”, $Q_s(u_1 \dots u_{N_s})$ – нелінійні хімічні реакції, N_s – кількість домішок. Нелінійні функції хімічних реакцій можуть бути представлені у вигляді

$$Q_s(u_1 \dots u_{N_s}) = - \sum_{i=1, i \neq s}^{N_s} \alpha_{si} u_i + \sum_{i=1, i \neq s}^{N_s} \sum_{j=1, j \neq s}^{N_s} \beta_{sij} u_i u_j, \quad s=1, \dots, N_s,$$

а функції джерел забруднення хімічними домішками – $f_s(t, z) = \sum_{j=1}^n q_{sj}(t) \delta(z - z^j)$,

де $s=1, \dots, N_s$, $q_{sj}(t)$ – інтенсивність джерел забруднення, n – кількість джерел.

В нерівності (1) $\psi_s \neq +\infty$ – це опуклий, обмежений знизу неперервно диференційований функціонал

$$\psi_s(v) = \int_{\Omega} \phi_s(v(z)) dz, \quad s=1, \dots, N_s,$$

Відповідна функція ϕ_s має такий вигляд:

$$\phi_s(v) \in \frac{d\psi_s}{dt}, \quad s=1, \dots, N_s.$$

У роботі досліджується процес самоочищення при перевищенні граничної концентрації забруднюючої домішки, який визначається системою вигляду

$$\psi_s = \begin{cases} \frac{1}{2} (m_s(t, z) u_s(t, z))^2, & u_s(t, z) \geq u_{s\max}(z), \\ 0, & u_s(t, z) < u_{s\max}(z) \end{cases}, \quad s=1, \dots, N_s,$$

де $m_s(t, z) \in M = L^\infty(\Theta)$ – коефіцієнт самоочищення (тут M – простір параметрів m_s з нормою $\|m_s\|_M = \|m_s\|_{L^\infty(\Theta)}$). Допустимі значення параметрів задаються виразом $M_{don} = \{m_s \in M \mid m_{s\max} \geq m_s \geq 0\}$.

Модель відома високими вимогами до обчислювальних ресурсів через велику кількість взаємодіючих домішок N_s , що призводить до великої розмірності СЛАР на етапі застосування скінчено-вимірної апроксимації.

Розв’язання системи варіаційних нерівностей (1) – (3) зводиться до розв’язання еквівалентної системи рівнянь:

$$\frac{\partial u_s}{\partial t} + A(\lambda)u_s + \phi(u_s) = f \text{ в } Q, \quad (4)$$

$$u_s|_{\Sigma} = 0, \quad (5)$$

$$u_s|_{t=0} = u_{s0} \text{ в } \Omega, \quad (6)$$

де $\phi(u_s)$ – невідома функція.

Рішенням системи (4) – (6) є пара $\{\hat{u}_s(t, z), \hat{m}(u; t, z)\}$, яку знайдено за умов забезпечення мінімуму функціоналу

$$J(m) = \sum_s \int_0^{t_k} \int_{\Omega} \left\{ \begin{array}{ll} (u_s - u_{s\max})^2, & u_s|_{\Gamma_B} \geq u_{s\max} \\ (m(u_s - u_{s\max}))^2, & u_s|_{\Gamma_B} < u_{s\max} \end{array} \right\} dz dt \rightarrow \inf_{m \in M_{don}}, \quad (7)$$

де m – коефіцієнт самоочищення.

Рішення системи рівнянь (4) – (6) базується на застосуванні методів скінчено-вимірної апроксимації і представлення їх у формі СЛАР.

Особливістю СЛАР є велика розмірність, суттєва розрідженість матриці спеціального виду, що веде до значних потреб оперативної пам’яті комп’ютерних систем. Частково подолати такого роду складності дозволяє запропоноване в відомих працях Х. Прайса та К. Коутса застосування процедур розв’язання СЛАР великої розмірності методом переміжних діагоналей (далі Д4). Разом з тим обчислювальна ефективність методу залишається недостатньою, що обумовлює подальші дослідження в цьому напрямі.

У роботі розроблено два методи обчислювальної реалізації СЛАР великої розмірності – прямий (на основі модифікованого Д4 та LU-декомпозиції, надалі, паралельний прямий МД4) та гібридний (на основі модифікованого Д4 та стабілізованого методу біспряжених градієнтів, надалі, паралельний гібридний МД4). Розроблені методи характеризуються низькими вимогами до оперативної пам’яті, що дозволяє ефективно використовувати їх у інформаційних технологіях екологічного моніторингу процесів забруднення атмосфери на паралельних обчислювальних системах з обмеженим обсягом пам’яті. Прикладом таких систем є кластери графічних прискорювачів, що відрізняються високою продуктивністю, масштабованістю та низькими вимогами до енергоспоживання.

На першому кроці організації паралельних методів (прямого та гібридного МД4) обчислювальної реалізації скінчено-вимірних вихідних рівнянь математичної фізики вирішується задача представлення цих рівнянь у формі розрідженої матричної моделі високої розмірності, яка є матричним аналогом СЛАР:

$$AU^{m+1} = B, m = \overline{0, t_k}, U^m|_{m=0} = U_0 \quad (8)$$

де A – розріджена матриця коефіцієнтів розмірності $(NP \times NP)$, B – відомий вектор-стовпець правих частин розмірності $(1 \times NP)$, $(0, t_k)$ – інтервал часу, m – номер кроку у часі.

Впорядковуючи невідомі матричної моделі (1) за схемою Д4, отримано матричне рівняння з блочною структурою матриці A :

$$\begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} \cdot \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}^{m+1} = \begin{bmatrix} B^* \\ F \end{bmatrix}, \quad (9)$$

де B^* та F – відомі праві частини, а структури підматриць A_1, A_2, A_3, A_4 , для випадку сітки 4×5 , наведено на рис. 1а (тут символом “*” позначено ненульові елементи підматриць).

Після виключення невідомих у нижній частині, рівняння (9) зводиться до вигляду:

$$\begin{bmatrix} A_1 & A_2 \\ 0 & A'_4 \end{bmatrix} \cdot \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}^{m+1} = \begin{bmatrix} B^* \\ F' \end{bmatrix}, \quad (10)$$

де структура підматриць A_1, A_2 та A'_4 , для випадку сітки 4×5 , наведена на рис. 1б.

Рівняння (10) можна вирішити для нижньої та верхньої частин окремо:

$$A'_4 U_2^{m+1} = F', \quad (11)$$

$$U_1^{m+1} = A_1^{-1} B^* - A_1^{-1} A_2 U_2^{m+1}, \quad (12)$$

Кожне з цих матричних рівнянь має розмір вдвічі менший ніж вихідне матричне рівняння (8). Результуюче рішення отримуємо у вигляді вектор-стовпця $U^{m+1} = [U_1, U_2]^{m+1}$.

В порівнянні з відомим методом Коперсмита-Вінограда зі складністю $O(n^{2.376})$, складність алгоритму на основі МД4 становить $9n + LSS\left(\frac{n}{2}\right)$, де $LSS(n)$ – складність розв’язання матричного рівняння розмірності n , $n = NP$.

Використання загальної задачі представлення скінченно-вимірних рівнянь математичної фізики у формі розріджених матричних рівнянь високої розмірності з впорядкування невідомих за схемою Д4 дозволяє організувати ефективну обчислювальну процедуру на базі систем з високим ступенем паралелізму таких, як графічні процесори та кластери на їх основі.

У таблиці 1. розглянуто порівняння витрат з розповсюдженими форматами зберігання матриць, окремо розглянуті два випадки розбиття області:

$[0, P] \times [0, P]$ та $[0, P] \times [0, 2P]$, $P=128$. Для кожного випадку знайдено кількість N_{float} ненульових елементів матриці A'_4 та накладні витрати для збереження позицій кожного ненульового елементу (N_{int}).

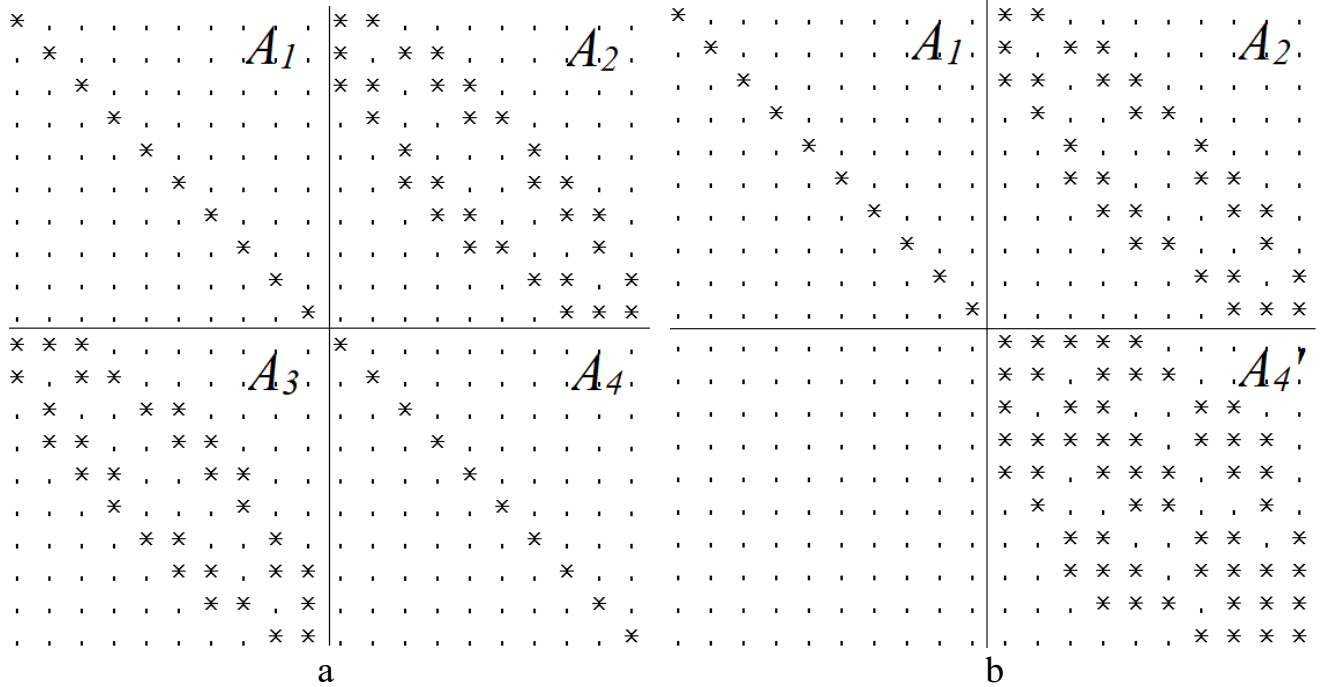


Рис. 1. Матриця A рівняння (9) модифікованого Д4 алгоритму для випадку $N=4$, $P=5$: (а) – до виключення невідомих у нижній частині; (б) – після виключення

У таблиці 1 використані наступні позначення форматів: DNS (dense) – зберігаються всі елементи матриці, включаючи нульові; CSR (compressed sparse row) – зберігаються координати рядків та ненульових елементів відносно початку рядка (без повторень рядків), COO (coordinate list) – зберігаються дві координати ненульового елементу, MD4 (modified D4) – координати ненульових елементів не зберігаються, обчислюються аналітично на етапі розв'язання матричних рівнянь.

Таблиця 1. Оцінки використання пам'яті для збереження A'_4 , $P=128$

Формат	N=P			N=2P		
	N_{float}	N_{int}	Всього, байт	N_{float}	N_{int}	Всього, байт
DNS	67108864	0	536870912	268435456	0	2147483648
CSR	72706	80898	905240	145922	162306	1816600
COO	72706	145412	1163296	145922	291844	2334752
MD4	72706	0	581648	145922	0	1167376

На другому кроці організації паралельних методів обчислювальної реалізації скінчено-вимірних вихідних рівнянь математичної фізики пропонується використовувати два розроблених паралельних методи (прямий та гібридний МД4).

Паралельний прямий метод МД4 заснований на застосуванні паралельного векторизованого блочного LU-розкладу з вибором провідного елемента для стрічкових систем LAPACK на етапі розв'язання матричного рівняння (11) та паралельної реалізації співвідношення (12), для систем зі спільною пам'яттю.

Матриця A'_4 розділюється на діагональні блоки, що не перетинаються, розмірності $(k'_u + n_b) \times (k_l + n_b)$, $k'_u = k_u + k_l$, де k_l – ширина нижньої частини рядка, k_u – верхньої, вигляду:

$$A_b^i = \begin{bmatrix} A_{11}^i & A_{12}^i & A_{13}^i \\ A_{21}^i & A_{22}^i & A_{23}^i \\ A_{31}^i & A_{32}^i & A_{33}^i \end{bmatrix},$$

де $A_{11}^i, A_{13}^i, A_{31}^i, A_{33}^i \in R^{n_b \times n_b}$, $A_{12}^i, A_{32}^i \in R^{n_b \times (k'_u - n_b)}$, $A_{21}^i, A_{23}^i \in R^{(k_l - n_b) \times n_b}$, A_{13}^i – верхня трикутна, A_{31}^i – нижня трикутна матриці, n_b – параметр алгоритму, i – індекс поточної ітерації.

Розрахунок LU-декомпозиції здійснюється із застосуванням реалізації BLAS (Basic Linear Algebra Subprograms – базові підпрограми лінійної алгебри), що забезпечує можливість перенесення програмної реалізації без змін вихідних кодів на нові моделі графічних процесорів. У термінах BLAS елементи паралельного розкладу $A_{ij} = L_{ij}U_{ij}$ обчислюються у такій послідовності:

1. Розраховується LU-розклад з перестановками (алгоритм Гауса з вибором провідного елемента) матриці: $(A_{11}, A_{12}, A_{13})^T$.
2. Обчислення DTRSM: $U_{12} \leftarrow L_{11}^{-1}A_{12}$.
3. Обчислення DGEMM: $A'_{22} \leftarrow A_{22} - L_{21}U_{12}$.
4. Обчислення DGEMM: $A'_{32} \leftarrow A_{32} - L_{31}U_{12}$.
5. Обчислення DTRSM: $U_{13} \leftarrow L_{11}^{-1}A_{13}$.
6. Обчислення DGEMM: $A'_{23} \leftarrow A_{23} - L_{21}U_{13}$.
7. Обчислення DGEMM: $A'_{33} \leftarrow A_{33} - L_{31}U_{13}$.
8. Обернена до кроку 1 перестановка системи: $(L_{11}, L_{12}, L_{13})^T$.

Метод МД4 для систем з розподіленою пам'яттю також може бути реалізований на базі використання технології SPIKE – декомпозиції бібліотеки Intel SpikePACK на етапі розв'язання матричного рівняння (11).

Обидва розглянуті варіанти першого паралельного прямого методу МД4 орієнтовані на використання графічних прискорювачів архітектури Nvidia CUDA. Низькі потреби оперативної пам'яті даних методів компенсують відповідний недолік графічних прискорювачів.

Другим з розроблених методів вирішення матричного рівняння (11) є паралельний гібридний метод МД4, заснований на використанні модифікованої Д4 декомпозиції та алгоритму біспряжених градієнтів. Запропонована у роботі покрокова процедура паралельної реалізації цього методу наведена нижче:

Крок 1. $x_0 = r_0^* = U_2^*$, де U_2^* – розв’язок на попередньому кроці, 0 для першого кроку;

$$r_0 = p_0 = F' - A'_4 x_0.$$

Крок 2. Цикл $j=0,1,\dots$ доки $|r_j| > \varepsilon$:

$$\alpha_j = \frac{(r_j, r_0^*)}{(A'_4 p_j, r_0^*)},$$

$$s_j = r_j - \alpha_j A'_4 p_j,$$

$$\omega_j = \frac{(A'_4 s_j, s_j)}{(A'_4 s_j, A'_4 s_j)},$$

$$x_{j+1} = x_j + \alpha_j p_j + \omega_j s_j,$$

$$r_{j+1} = s_j - \omega_j A'_4 s_j,$$

$$\beta_j = \frac{\alpha_j (r_{j+1}, r_0^*)}{\omega_j (r_j, r_0^*)},$$

$$p_{j+1} = r_{j+1} + \beta_j (p_j - \omega_j A'_4 p_j).$$

Крок 3. $U_2 = x_{j+1}$, де матричні та векторні операції обчислюються паралельно.

Останній метод пристосовано для застосування кластерів графічних процесорів, в тому числі реалізовано алгоритм біспряжених градієнтів для кластерів з підтримкою CUDA та MPI.

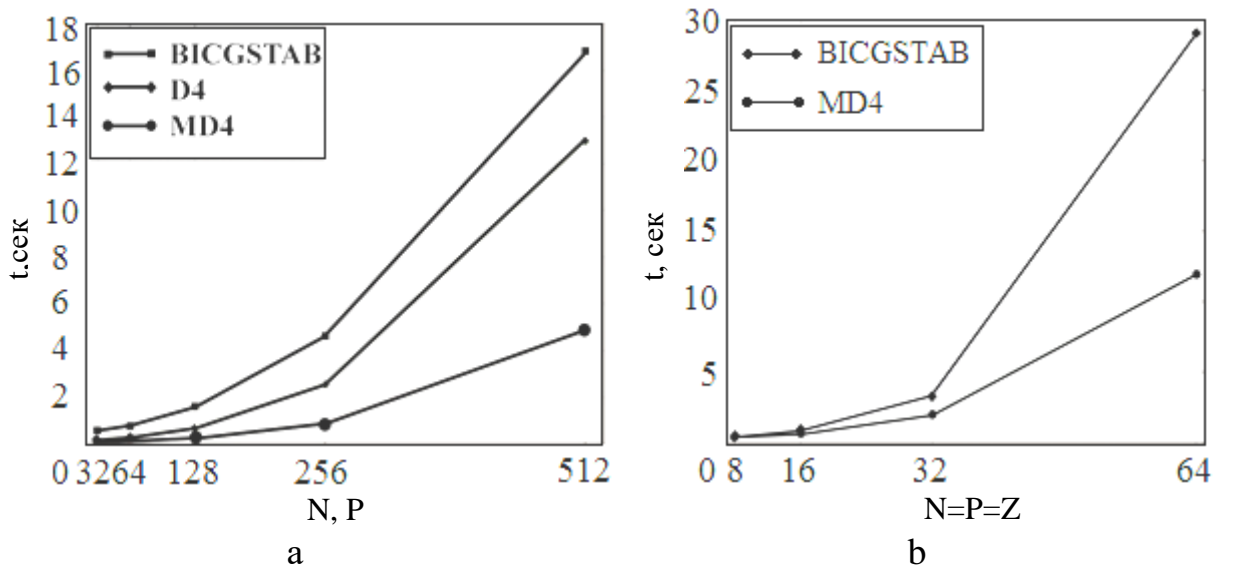


Рис. 2. Порівняння паралельних прямого та гібридного МД4 методів зі стандартним стабілізованим методом біспряжених градієнтів (BiCGSTAB) для дво- (а) та тривимірних (б) задач на графічному процесорі Nvidia GeForce 9600M GT

Працездатність та обчислювальна ефективність програмної реалізації розроблених методів у порівнянні зі стандартним для дво- та тривимірних задач перевірялась шляхом обчислювального експерименту на графічному процесорі CUDA (рис. 2). Запропоновані методи демонструють вищу швидкість у порівнянні з класичним D4 та ітераційним методом біспряжених градієнтів.

В **третьому розділі** удосконалено метод параметричної ідентифікації функцій джерел моделі, який відрізняється врахуванням ефектів хімічної взаємодії домішок та аномальних ефектів самоочищення атмосфери. Метод характеризується підвищеною швидкістю за рахунок використання розроблених в другому розділі паралельних методів вирішення СЛАР великої розмірності та використання графічних прискорювачів архітектури CUDA.

У випадку відсутності аномальних властивостей самоочищення модель (1) – (3) може бути зведена до відомої моделі UNI-DEM (Unified Danish Eulerian Model):

$$\frac{\partial u_s}{\partial t} = -\frac{\partial(c_i u_s)}{\partial z_i} + \frac{\partial}{\partial z_i} \left(k_i(z) \frac{\partial u_s}{\partial z_i} \right) + f_s(t, z) + \quad (13)$$

$$+ Q_s(u_1 \dots u_{N_s}) - (\kappa_{1s} + \kappa_{2s}) u_s, \quad s = 1, \dots, N_s, i = 1, 2, 3$$

$$\int_{\Omega} d(z) u \xi dz = Q_s(u_1 \dots u_{N_s}) - (\kappa_{1s} + \kappa_{2s}) u_s, \quad s = 1, \dots, N_s,$$

$$f_s(t, z) = \sum_{j=1}^n q_{sj}(t) \delta(z - z^j) \quad (14)$$

де $f_s(t, z)$ – функції джерел забруднення атмосфери хімічними домішками, u_s – концентрація домішки s , c_i – швидкість вітру, $k_i(z)$ – коефіцієнти дифузії, $q_{sj}(t)$ – інтенсивність джерел, κ_{1s} , κ_{2s} – коефіцієнти осадження, $Q_s(u_1 \dots u_{N_s})$ – нелінійні хімічні реакції, $N_s = \{35, 56, 168\}$ – кількість домішок.

Нелінійні функції хімічних реакцій мають вигляд

$$Q_s(u_1 \dots u_{N_s}) = -\sum_{i=1}^{N_s} \alpha_{si} u_i + \sum_{i=1}^{N_s} \sum_{j=1}^{N_s} \beta_{sij} u_i u_j, \quad s = 1, \dots, N_s. \quad (15)$$

Початкові та граничні умови мають вигляд

$$\begin{aligned} u_s|_{t=0} &= u_{s_0}, \quad s = 1, \dots, N_s, \\ u_s|_{\Gamma} &= u_{s_{\Gamma}}, \quad s = 1, \dots, N_s. \end{aligned} \quad (16)$$

Модель розповсюдження забруднюючих домішок в атмосфері у формі системи варіаційних нерівностей (1) – (3), процедура обчислювального розв'язання якої записана у формі систем рівнянь (4) – (7), враховуючи ефекти самоочищення атмосфери та хімічної взаємодії домішок N_s , вимагає великої кількості обчислювальних ресурсів. Для обчислювальної реалізації вихідні модельні співвідношення переводяться у скінчено-вимірну форму та представляються у вигляді матричної моделі (8). Останнє дозволяє використовувати розроблені в розділі 2 паралельні прямий та гібридний методи МД4.

У розділі запропоновано метод розв'язання задачі ідентифікації функцій джерел процесів розповсюдження забруднюючих домішок в атмосфері, представлених у формі розширеної моделі UNI-DEM (співвідношення (13) – (16)), пряме вимірювання, яких не є можливим. Актуальність цієї задачі визначається широким розповсюдженням у екологічній практиці.

Аналіз задачі ідентифікації функції джерел розглянуто в такій постановці. На першому етапі розглядається процес, який описується рівняннями (13) – (16), необхідно за вимірами y_s^{sum} функції $u_s(t, z^i)$ в точках $z^1, \dots, z^n \in \Omega$ отримати оцінки $f_s(t, z)$ (рівняння (14)), які б мінімізували функціонал

$$J(u_s, f_s) = \int_0^{t_k} \sum_{i=1}^n [y_{is}^{sum}(t) - y_{is}(t)]^2 dt \rightarrow \inf_{f \in F_{don}}, \quad (17)$$

де $y_s(t, z^i) = h(t, z^i)u_s(t, z^i)$ (тут $h(t, z^i)$ – відомий оператор), F_{don} – множина допустимих забруднюючих функцій процесу.

Для вирішення задачі пошуку екстремуму функціоналу $J(u_s, f_s)$ рівняння (17) за умови виконання системи рівнянь (13) – (16) застосовується принципу Лагранжа.

Пошук сукупності невідомих параметрів джерел забруднення атмосфери хімічними домішками виконується за допомогою градієнтної процедури

$$f_s^{i+1} = \text{Pr} \left\{ f_s^i - \omega_f \left(\frac{\partial L(u_s, p_s, f_s)}{\partial f_s} \right)^i \right\},$$

де $p_s(t, z)$ – спряжена змінна, $L(u_s, p_s, f_s)$ – лагранжіан, i – індекс поточної ітерації, f^0, ω_f – задані.

Для отримання числового рішення до математичної моделі (4) – (7) застосовується процедура скінчено-вимірної апроксимації з подальшим представленням її у формі СЛАР великої розмірності (матрична модель (8)). Матрична модель, в свою чергу, вирішується із застосуванням паралельних методів (прямого та гібридного МД4).

Представлений метод параметричної ідентифікації функцій джерел забруднення в моделях аномальних процесів забруднення атмосфери з врахуванням процесів самоочищення та використання графічних процесорів Nvidia CUDA був досліджений шляхом проведення обчислювальних експериментів. Зокрема, вирішувалась задача ідентифікації інтенсивності стаціонарних джерел забруднення атмосфери викидами p_1 – діоксиду сірки, p_2 – фенолу та p_3 – аміаку промислового об'єкту за даними наданими ТОВ «Укрекотехінвест». Розв'язки задачі порівнювалися з отриманими інструментально прямими вимірами відповідних викидів $p_1^* = 0.28$ мг/сек, $p_2^* = 0.0464$ мг/сек, $p_3^* = 0.1644$ мг/сек (рис.3а). Результати експериментів свідчать про високі обчислювальні характеристики запропонованого методу, зокрема, похибка отриманих результатів не перевищує 7%, а час виконання процедури ідентифікації з використанням графічних процесорів Nvidia CUDA

скоротився в 10-12 разів в порівнянні з застосуванням стандартних суперкомп'ютерних технологій (рис.3b).

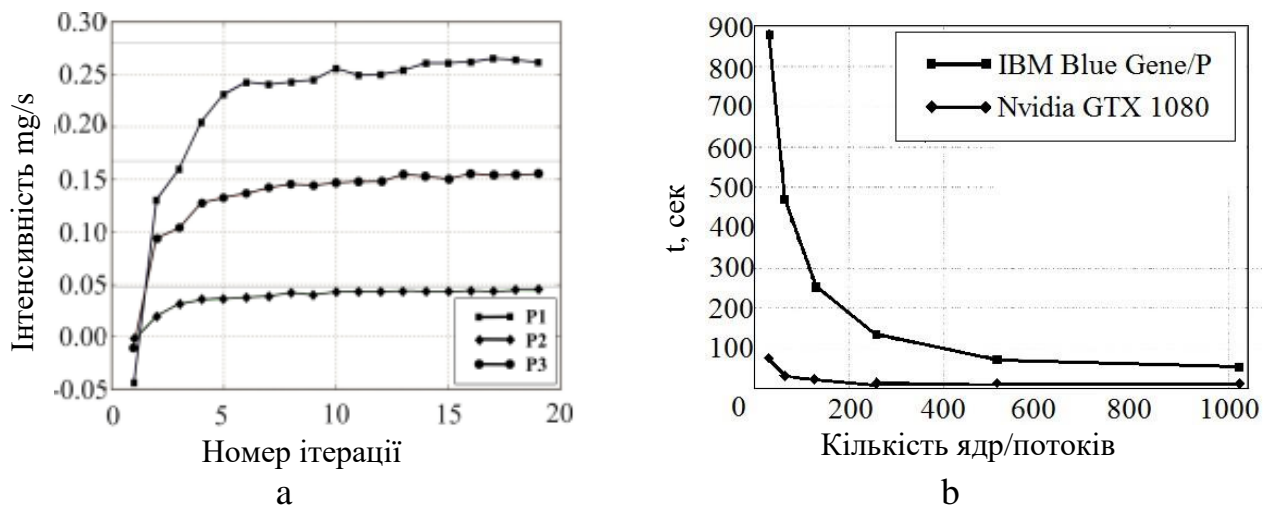


Рис.3. Порівняльні характеристики вирішення задачі ідентифікації інтенсивності стаціонарних джерел забруднення атмосфери: (а) – порівняння отриманих результатів з прямими вимірами відповідних викидів, (б) – порівняння часу виконання процедури ідентифікації із застосуванням графічних процесорів Nvidia CUDA та суперкомп'ютеру IBM Blue Gene/P

В **четвертому розділі** розроблено інформаційну технологію моніторингу аномальних процесів забруднення атмосфери, яка складається з поєднання технологій збору, обробки, накопичення та надання кінцевому користувачу результатів моніторингу. Важливою складовою цієї технології є спеціалізована комп'ютерна система на базі графічних прискорювачів архітектури Nvidia CUDA, яка характеризується підвищеною швидкістю, низькою вартістю елементної бази на відміну від традиційних суперкомп'ютерних технологій.

Відомим недоліком комп'ютерних систем, що базуються на графічних прискорювачах, є обмеження на обсяг оперативної пам'яті, яка надається користувачам. Для подолання цієї проблеми, в запропонованій комп'ютерній системі, використовується розроблені в дисертації паралельні обчислювальні методи моделювання процесів забруднення атмосфери, які відрізняються низькими вимогами до оперативної пам'яті на відміну від традиційних.

Запропоновану інформаційну технологію закладено в основу створеного спеціалізованого програмно-апаратного комплексу "Моніторинг". Комплекс складається з сенсорів, системи агрегації даних, системи збереження, геопросторової бази даних (БД), системи представлення геопросторової інформації, системи візуалізації, системи керування, системи організації обчислень та обчислювальної інфраструктури, в якості якої можуть бути використані гібридний кластер з графічними прискорювачами, персональний суперкомп'ютер, традиційний кластер або грід-середовище (рис. 4, 5).

В якості обчислювальної платформи для сенсорів використано мікроконтролер ESP8266. Вибір обумовлений такими перевагами: низька вартість; підтримка стеку TCP/IP; вбудований WiFi; можливість розроблення програмного забезпечення на

мові Python (реалізація MicroPython). Для передавання даних спостережень, зокрема WiFi, доступні модулі розширення з підтримкою ZigBee та GSM при необхідності мобільного застосування – GPS. Розроблене програмне забезпечення сенсору здійснює періодичне вимірювання стану середовища, збереження результатів у вбудованій флеш-пам'яті та передавання в систему агрегації даних. При передаванні використовуються протоколи HTTPS (модуль MicroPython urequests) дані в форматі json (модуль ujson).

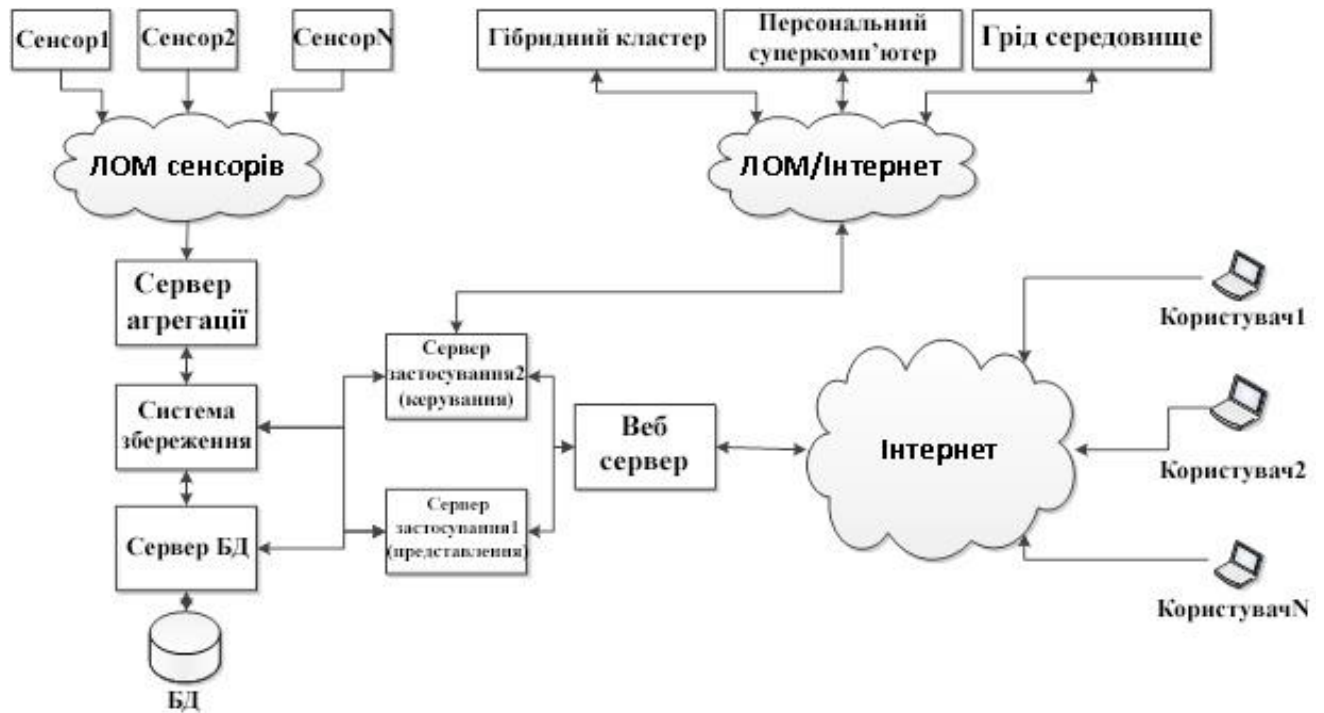


Рис. 4. Архітектура розробленої технології

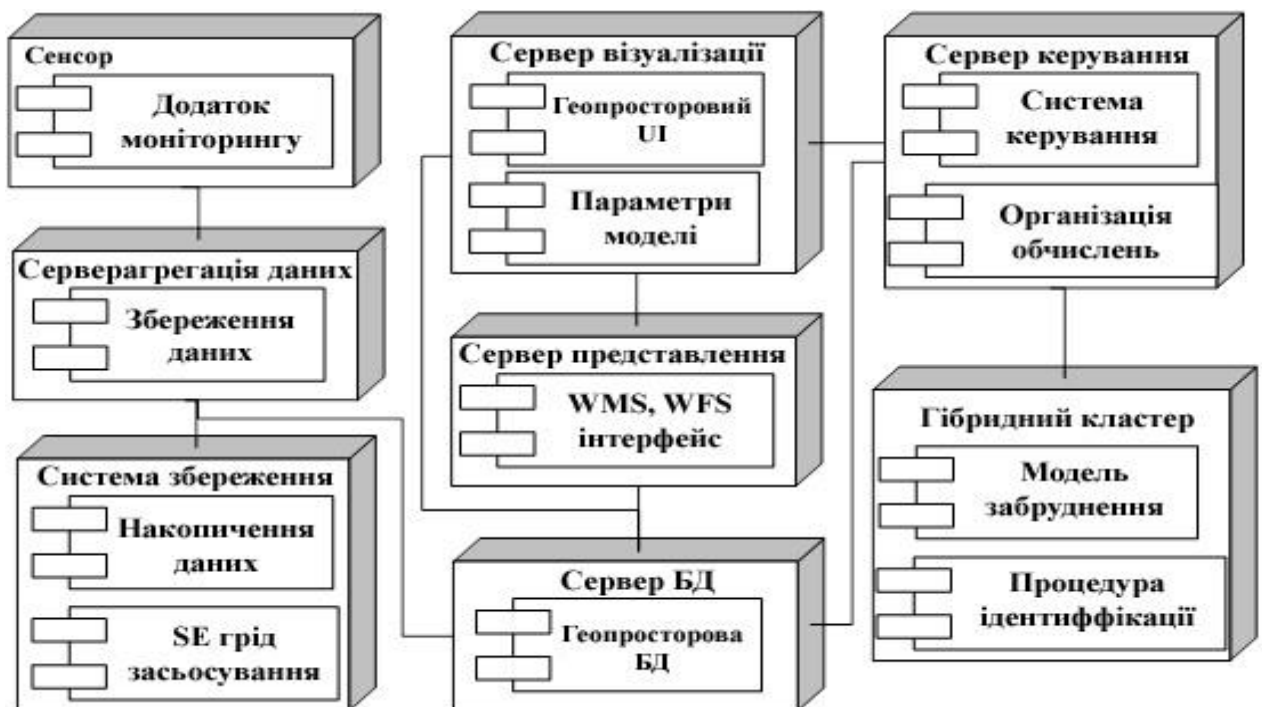


Рис. 5. Діаграма розгортання системи

Система агрегації забезпечує збір даних сенсорів, заповнення геопросторової БД та системи збереження. При програмній реалізації використано Python, мікроплатформу Flask (web-сервер прийому даних сенсорів), SQLAlchemy (взаємодія з БД), модулі netCDF4-python, pyKML (збереження в форматах netCDF та KML відповідно). При взаємодії з сервером БД використовується SQL, система збереження підключена до локальної файлової системи.

Для реалізації системи збереження використовується мережеве сховище (за протоколом NFS) та Amazon S3 (s3fs, в залежності в обсягу оброблюваних даних). При використанні ґрід в якості обчислювальної інфраструктури, зовнішній доступ до системи збереження організовано за протоколом GridFTP.

Геопросторові дані сенсорів, історичні дані, інформація про параметри досліджуваної області, фізичні характеристики забруднюючих домішок, результати моделювання зберігаються у БД, СУБД PostgreSQL з розширенням PostGIS. Завдяки цьому забезпечується прозоре оброблення геопросторових даних на рівні БД та можливість абстрагування досліджуваних моделей від форматів просторових даних.

Система представлення геопросторових даних забезпечує доступ до вихідних даних та результатів моделювання за протоколами OGC WMS та WFS. Розділення функцій збереження, представлення та візуалізації даних, використання стандартизованих протоколів OGC спрощує інтеграцію запропонованої системи з існуючими системами підтримки прийняття рішень та робить можливим прямий доступ до результатів моніторингу з популярних геоінформаційних систем (таких як ESRI ArcGIS, QGIS та інші). Для програмної реалізації використано MapServer, дані отримуються з описаної вище БД.

Завдяки тому, що система візуалізації реалізована у вигляді web-сервісу, для доступу до неї достатньо web-браузеру з підтримкою HTML5. Висока продуктивність при візуалізації просторової інформації на стороні клієнта досягається завдяки використанню бібліотеки OpenLayers, що інтегрує результати моніторингу і моделювання з системи представлення, супутникові карти з Google Maps, дані спостережень інших організацій. Крім візуалізації в системі можуть бути задані параметри моніторингу, та запущений процес моделювання. Відповідний web-додаток реалізовано на базі Python Flask та інтегровано за допомогою mod_wsgi з web-сервером Apache.

Система керування отримує параметри моделювання з системи візуалізації, вхідні дані з БД та забезпечує формування обчислювальних завдань для системи організації обчислень. Підтримуються специфікації завдань для систем керування навантаженням кластеру SLURM, Torque, ґрід-систем на базі ARC, gLite, UMD, та прямий запуск програмної моделі шляхом виконання mpirun, ssh/rexec. Сформоване завдання відправляється до системи організації обчислень, після закінчення моделювання результати заносяться в БД. Для взаємодії системи керування та системи організації обчислень використовується ZeroMQ.

Система організації обчислень забезпечує запуск на виконання обчислювальних завдань, моніторинг виконання та отримання результатів моделювання. Взаємодія з обчислювальними системами здійснюється шляхом прямого запуску утиліт командного рядка систем керування навантаженням

кластера та проміжного програмного забезпечення ґрид. При програмній реалізації модуля менеджера завдань використано Python multiprocessing, PyZMQ.

Програмна реалізація моделі забруднення атмосфери та методів параметричної ідентифікації мовою C++ із застосуванням MPI та розширень компілятора Nvidia CUDA визначає вимоги до обчислювальних елементів системи. Необхідна наявність CUDA toolkit в залежності від системи OpenMPI (гібридний кластер), підтримка OpenMP (персональний суперкомп'ютер). Розділення системи організації обчислень та обчислювальної інфраструктури дозволяє абстрагувати систему керування обчислювальною інфраструктурою – програмна реалізація моделі не потребує адаптації при перенесенні з гібридного кластера у ґрид або хмарне середовище.

Розроблена інформаційна технологія екологічного моніторингу була реалізована у спеціалізованому програмно-апаратному комплексі «Моніторинг». Комплекс реалізовано на гібридній високопродуктивній обчислювальній системі архітектури Nvidia CUDA. Останнє забезпечує підвищення оперативності і точності моніторингу та розширення області його застосування за рахунок можливості використання мобільних комплексів. Розроблений комплекс використовується ТОВ «Укрехотехінвест» для вирішення практичних задач екологічного моніторингу та в навчальному процесі Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського».

ВИСНОВКИ

У дисертаційній роботі розв'язано актуальну науково-практичну задачу створення інформаційної технології екологічного моніторингу процесів забруднення атмосфери з використанням високопродуктивних обчислень на базі графічних процесорів загального призначення. Проведені дослідження дозволили отримати нові наукові результати, які мають істотні переваги перед існуючими, а саме:

1. На основі аналізу інформаційних технологій екологічного моніторингу зроблено висновок про актуальність задачі розроблення інформаційної технології моніторингу аномальних процесів забруднення атмосфери хімічно-взаємодіючими домішками, які базуються на спеціалізованих математичних методах та дають можливість застосувати високопродуктивні графічні прискорювачі загального призначення, що веде до підвищення обчислювальної ефективності моніторингу.

2. Розроблено метод паралельної обчислювальної реалізації моделей, використовуваних при дослідженні процесів забруднення атмосфери, який характеризується можливістю ефективного застосування гібридних обчислювальних систем на базі графічних прискорювачів архітектури Nvidia CUDA, та ефективним використанням модифікованих методів декомпозиції МД4 (паралельний гібридний МД4 та паралельний прямий МД4) вирішення СЛАР високої розмірності та використовує, в порівнянні з іншими, на 30% менше оперативної пам'яті.

3. Удосконалено метод параметричної ідентифікації функцій джерел в моделях процесів забруднення атмосфери хімічно-взаємодіючими домішками з врахуванням аномальних ефектів самоочищення на основі моделі UNI-DEM. З використанням графічних процесорів Nvidia CUDA метод дозволяє скоротити час виконання

процедури в 10-12 разів у порівнянні з стандартними суперкомп'ютерними технологіями, при цьому похибка отриманих результатів не перевищує 7%.

4. Вперше розроблено інформаційну технологію моніторингу аномальних процесів забруднення атмосфери, яку складає поєднання технологій збору, оброблення, накопичення та надання кінцевому користувачу результатів моніторингу, на основі методів реалізації моделей забруднення атмосфери, що уможливають використання гібридних високопродуктивних обчислювальних систем архітектури Nvidia CUDA, і забезпечують підвищення оперативності і точності моніторингу та розширення області його охоплення за рахунок можливості використання мобільних комплексів.

5. Створено спеціалізований програмно-апаратний комплекс “Моніторинг”, який реалізує запропоновану інформаційну технологію моніторингу процесів забруднення атмосфери та характеризується меншими вимогами до оперативної пам'яті, що дозволяє використовувати гібридні паралельні обчислювальні системи архітектури Nvidia CUDA з підвищеною оперативністю стаціонарних або мобільних комплексів моніторингу.

6. Розроблена інформаційна технологія використовується в ТОВ “Укрекотехінвест” і навчальному процесі кафедр інженерної екології Інституту енергозбереження та енергоменеджменту та інформаційної безпеки Фізико-технічного інституту Національного технічного університету України “Київський політехнічний інститут імені Ігоря Сікорського”.

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

1. Ільїн М.І., Новіков О.М. Ідентифікація інтенсивності джерел забруднення атмосфери на базі гібридних обчислювальних систем // Системні дослідження та інформаційні технології. – 2017. – 3. – С. 21-29. (Проіндексовано в міжнародних наукометричних базах: Index Copernicus, DOAJ, РІНЦ). *Здобувачу належить розроблений алгоритм та вдосконалення методу ідентифікації інтенсивності джерел.*

2. Ильин Н.И. Информационная технология мониторинга загрязнения атмосферы химически взаимодействующими примесями с учетом их аномальных свойств // Управляющие системы и машины. – 2017. – 4. – С. 59-67. (Проіндексовано в міжнародній наукометричній базі РІНЦ).

3. Ільїн М.І. Модифікація методу обчислювальної реалізації крайових задач на основі Д4 декомпозиції для гібридних систем на базі графічних процесорів // Вісник НТУУ КПІ: Інформатика, управління та обчислювальна техніка. – 2017. – 65. – С. 63-67. (Проіндексовано в міжнародній наукометричній базі РІНЦ).

4. Новіков О.М., Терещенко І.М., Ільїн М.І., Семенов О.В. Дослідження методів паралельних обчислень для подолання трансобчислювальної складності при розв'язуванні екологічних проблем навколишнього середовища // Наукові вісті КПІ. – 2011. – 1. – С. 79-83. *Здобувачу належить реалізація та експериментальне дослідження методів в ґрид-середовищі Українського академічного ґрид.*

5. Парусимов Г.В., Ильин Н.И., Рожко Р.В. Разработка и тестирование параллельного алгоритма для решения краевой задачи уравнения параболического

типа в УАГ // Кибернетика и вычислительная техника. – 2010. – 161. – С. 15-27. *Здобувачу належить реалізація та експериментальне дослідження методів в ґрид-середовищі Українського академічного ґрид.*

6. Ильин Н.И. Модель производительности системы визуализации спутниковых данных // Управляющие системы и машины. – 2009. – 4. – С. 15-22.

7. Ильин, Н.И. Кикоть Н.И. Реализация модифицированного алгоритма Д4 декомпозиции в multiGPU системах / Тези за матеріалами міжнародної наукової конференції студентів, аспірантів і молодих вчених «Теоретичні і прикладні аспекти кібернетики». – Київ: КНУ, 21-25 лютого 2011 р. – С. 155-157. *Здобувачу належить алгоритм для multiGPU систем.*

8. Ільїн М.І., Новіков О.М., Кікоть М.І. Моделювання процесів розповсюдження шкідливих домішок в атмосфері для задач екологічного прогнозування в умовах надзвичайних ситуацій / Матеріали VI міжнародної конференції “Екологічна безпека: проблеми і шляхи вирішення”. – Харків: УкрНДІЕП. – 2010. – С. 105-107. *Здобувачу належить розроблення моделі та експериментальне дослідження.*

9. Новиков А.Н., Ильин Н.И., Кикоть Н.И. Применение GPU кластеров в задаче идентификации функций источников загрязнений в модели UNI-DEM / Матеріали міжнародної науково-технічної конференції “Системний аналіз та інформаційні технології”. – Київ: НТУУ КПІ. – 2011. – С. 477-478. *Здобувачу належить розроблення методів і алгоритмів для GPU кластерів.*

АНОТАЦІЯ

Ільїн М.І. Інформаційна технологія екологічного моніторингу процесів забруднення атмосфери з використанням високопродуктивних обчислень. На правах рукопису.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.06 – інформаційні технології. Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», МОН України, Київ, 2019.

В дисертаційній роботі розроблено метод паралельної обчислювальної реалізації моделей, що виникають при дослідженні процесів забруднення атмосфери, який відрізняється можливістю ефективного застосування гібридних обчислювальних систем на базі графічних прискорювачів архітектури Nvidia CUDA і базується на використанні модифікованого методу Д4 декомпозиції, що дозволяє зменшити на 30% використання оперативної пам'яті. Вдосконалено метод параметричної ідентифікації функцій джерел забруднення в моделях аномальних процесів забруднення атмосфери на основі моделі UNI-DEM з врахуванням аномальних процесів самоочищення. Вдосконалений метод відрізняється підвищеною швидкістю за рахунок використання розроблених прямого та гібридного паралельних методів вирішення СЛАР великої розмірності та дозволяє скоротити час моделювання в 10-12 разів в порівнянні з реалізацією без застосування GPGPU. Розроблено інформаційну технологію моніторингу аномальних процесів забруднення атмосфери хімічно-взаємодіючими домішками,

що базується на гібридних високопродуктивних обчислювальних системах архітектури Nvidia CUDA, та відрізняється підвищеною ефективністю і меншими вимогами до оперативної пам'яті. Для використання інформації технології розроблено програмно-апаратний комплекс "Моніторинг". Програмно-апаратний комплекс складається з сенсорів, системи агрегації даних, системи збереження, геопросторової бази даних (БД), системи представлення геопросторової інформації, системи візуалізації, системи керування, системи організації обчислень та обчислювальної інфраструктури в якості якої можуть бути використані гібридний кластер з графічними прискорювачами, персональний суперкомп'ютер, традиційний кластер або грід середовище.

Розроблену інформаційну технологію було використано для вирішення практичних задач ТОВ "Укрекотехінвест" та в навчальному процесі Національного технічного університету України "Київський політехнічний інститут імені Ігоря Сікорського".

Ключові слова: екологічний моніторинг, односторонні фізичні процеси, параметрична ідентифікація, моделі навколишнього середовища, кластер графічних процесорів.

АННОТАЦИЯ

Ильин Н.И. Информационная технология экологического мониторинга процессов загрязнения атмосферы с использованием высокопродуктивных вычислений. На правах рукописи.

Диссертация на соискание ученой степени кандидата технических наук по специальности 05.13.06 – информационные технологии. Национальный технический университет Украины "Киевский политехнический институт имени Игоря Сикорского", МОН Украины, Киев, 2019.

В диссертационной работе разработан метод параллельной вычислительной реализации моделей, возникающих при исследовании процессов загрязнения атмосферы, который отличается возможностью эффективного применения гибридных вычислительных систем на базе графических ускорителей архитектуры Nvidia CUDA. Метод основан на использовании модифицированного метода Д4 декомпозиции и использует на 30% меньше оперативной памяти. Усовершенствован метод параметрической идентификации функций источников загрязнения в моделях аномальных процессов загрязнения атмосферы на основе модели UNI-DEM с учетом аномальных процессов самоочищения. Метод отличается повышенным быстродействием за счет использования разработанных прямого и гибридных параллельных методов решения СЛАУ большой размерности и позволяет сократить время моделирования в 10-12 раз по сравнению с реализацией без применения GPGPU. Разработана информационная технология мониторинга аномальных процессов загрязнения атмосферы химически взаимодействующими примесями, основанная на гибридных высокопроизводительных вычислительных системах архитектуры Nvidia CUDA, и отличается повышенной эффективностью и меньшими требованиями к оперативной памяти. Для использования информации технологии

разработан программно-аппаратный комплекс "Мониторинг". Программно-аппаратный комплекс состоит из сенсоров, системы агрегации данных, системы хранения, геопространственной базы данных (БД), системы представления геопространственной информации, системы визуализации, системы управления, системы организации вычислений и вычислительной инфраструктуры в качестве которой могут быть использованы гибридный кластер с графическими ускорителями, персональный суперкомпьютер, традиционный кластер или грид среда.

Предложенная информационная технология была использована для решения практических задач ООО "Укрехотехинвест" и в учебном процессе Национального технического университета Украины "Киевский политехнический институт имени Игоря Сикорского".

Ключевые слова: экологический мониторинг, односторонние физические процессы, параметрическая идентификация, модель окружающей среды, кластер графических процессоров.

ABSTRACT

Ilin M.I. Information technology of ecological monitoring for high performance computing systems. As a manuscript.

A Philosophy Doctor of Technical Sciences thesis in specialty 05.13.06 — information technologies. National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ministry of Education and Science of Ukraine, Kyiv, 2019.

In the dissertation work the method of parallel computational implementation of models that arise in the atmospheric pollution processes is developed. The method differs in the ability to effectively use hybrid computing systems based on graphic accelerators of the Nvidia CUDA architecture based on the use of the modified D4 decomposition method and uses 30% less RAM. The method of parametric identification of the sources of pollution functions in models of unilateral atmospheric pollution processes based on the UNI-DEM model is proposed. Method allows taking into account abnormal self-cleaning processes. Developed method is distinguished by high speed due to the use of developed direct and hybrid parallel methods of solving large-scale SLAE and allows shortening the simulation time in a 10-12 times compared to the implementation without the use of GPGPU. The information technology of monitoring the unilateral processes of atmospheric pollution by chemically interacting impurities, based on Nvidia CUDA hybrid high performance computing systems, is developed. The developed technology is characterized by increased efficiency and lower requirements to RAM.

The dissertation consists of an introduction, four sections, conclusions and list of sources used.

The introduction substantiates the relevance of the topic of the dissertation, formulates the purpose and tasks of the research, the scientific novelty and the practical significance of the results. Data on the implementation of the results of work, its testing, publications and personal contribution of the applicant are given.

In the first section a critical analysis of the current state of simulation of the process of spreading pollutants in the atmosphere is carried out. A review of existing models and methods of atmospheric diffusion processes research has been performed. The recommendations on the use of harmful contaminant propagation models have been analyzed by the Environmental Protection Agency Support Center for Regulatory Atmospheric Modeling and the European Environment Agency Model Documentation System. On the basis of the analysis, the relevance of the topic of the dissertation study due to the computational complexity of the Euler models of the distribution of harmful impurities in the atmosphere and the optimization problem of the parametric identification of the one-sided diffusion-transfer process are limited to the limiting concentration of the contaminating impurity. The actuality of creation of high-performance realizations of SLAE algorithms and associated information technologies on the basis of clusters of graphics processors of architecture of Nvidia CUDA is shown.

The second section is devoted to solving the problem of increasing the computational efficiency of simulation of atmospheric pollution by chemically interacting impurities, taking into account abnormal properties. The solution is based on the use of methods of finite-dimensional approximation of the initial equations of mathematical physics with the subsequent representation of them in the form of a system of linear algebraic equations (hereinafter SLAE). The peculiarity of these equations is the large dimensionality and the substantial variety of specialty. Part of the overcoming of this kind of complexity is proposed in the works of H. Price and C. Coats on the basis of the application of procedures for solving large-scale SLAEs using the method with the preceding diagonals (hereinafter D4) for arranging the nodes of the spatial domain. Two modifications of the method of computing implementation are suggested in the work – parallel hybrid (based on modified D4 and biconjugate gradient stabilized method, hybrid MD4) and direct (based on MD4 and LU decomposition, direct MD4). The modifications of the method differ in low requirements for RAM, which allows them to be effectively used in information technology on high-parallel computing systems with a limited amount of memory. An example of such systems is a cluster of graphics processors that are characterized by high performance, scalability and high energy efficiency.

In the third section the model of atmospheric pollution processes by chemical interacting impurities is proposed taking into account the abnormal properties of self-purification of the atmosphere when the limiting concentration of the contaminating impurity reaches and the task of parametric identification of sources of polluting impurities of the model is solved. The section proposes a method for solving the parametric identification of the function of sources of the extended model UNI-DEM. Compared with the published results of UNI-DEM, the proposed methods allow to shorten simulation time by 10-12 times.

In the fourth section an information technology for monitoring the processes of atmospheric pollution by chemical interacting impurities is developed, taking into account abnormal properties. To use information technology software and hardware complex "Monitoring" is developed. Developed complex consists of sensors, data aggregation systems, storage systems, geospatial database (DB), systems for representing geospatial information, visualization systems, control systems, computing systems and computing infrastructure, which can be used as a hybrid cluster with graphic accelerators, personal

supercomputer, traditional cluster or grid environment. In order to assess the accuracy of the developed methods and the model of section 3, the software and hardware complex was used as an auxiliary vehicle in assessing the state of pollution of the atmosphere by stationary sources of industrial emissions and in the construction of pollution maps on the basis of LLC ``Ukrekotekhinvest". As a result of the model verification, the deviation was less than 7%, which suggests the possibility of practical application of the developed methods as an auxiliary means in assessing pollutant emissions in the enterprise, in particular, pollution of the atmosphere from stationary sources of industrial emissions. The proposed information technology was used to solve practical problems of ``Ukrekotekhinvest LLC" and in the educational process of the Physics and Technology Institute of the National Technical University of Ukraine "Igor Sikorsky Polytechnic Institute".

Keywords: ecological monitoring, unilateral physical processes, parametric identification, environmental model, GPU clusters.